Applications for *qg* Tagging

Jennifer Thompson

Universität Heidelberg

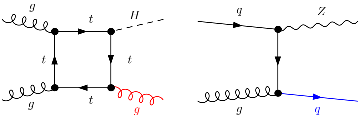
15.11.2018

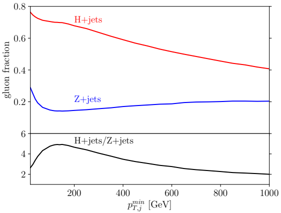
Gregor Kasieczka[1], Nicholas Kiefer[2], Tilman Plehn[2], Michael Russell[2]
Jennifer Thompson[1]

[1] Institut für Experimentalphysik Universität Hamburg [2] ITP Universität Heidelberg

Two Benchmarks

$H \to$ invisible+jet





$\longrightarrow$ Signal purity depends on p_T

$Z' \to q\bar{q}$

$$\mathcal{L} \supset g_q \sum_q Z' \bar{q} \gamma^\mu \gamma^5 q$$

- Pure quark signal
- Large QCD background
  $\longrightarrow$ Mostly gluon jets
- Leading 2 jets important

Lola: Machine Learning with jet consituents

- CoLa (Combination Layer) includes (trainable) 4-momenta
  $\longrightarrow$ Simulate a jet algorithm
- LoLa (Lorentz Layer) transforms includes physical parameters
  $\longrightarrow$ e.g. m, p_T
- Already applied to top tagging and autoencoders
  A. Butter, G. Kasieczka, T. Plehn, M. Russell arXiv:1707.08966 [hep-ph]

  T. Heimel, G. Kasieczka, T. Plehn, **J. Thompson** arXiv:1808.08979 [hep-ph]
- Now consider another interesting problem
  - **quark gluon discrimination**
    P. Komiske, E. Metodiev, M. Schwartz arXiv:1612.01551 [hep-ph]

    E. Metodiev, J. Thaler arXiv:1802.00008 [hep-ph]
  - hard process dependent
  - No strict definition from QCD

Cola and LoLa in Equations

The combination layer (CoLa) acts on 4-momenta $k_{\mu,i}$:

$$k_{\mu,i} \xrightarrow{\text{CoLa}} \widetilde{k}_{\mu,j} = k_{\mu,i} \; C_{ij}$$

and Lorentz layer (LoLa)

$$\tilde{k}_j \xrightarrow{\text{LoLa}} \hat{k}_j = \begin{pmatrix} m^2(\tilde{k}_j) \\ p_T(\tilde{k}_j) \\ w_{jm}^{(E)} \, E(\tilde{k}_m) \\ w_{jm}^{(d)} \, d_{jm}^2 \end{pmatrix}$$

$\longrightarrow$ make d_{jm}^2 trainable:

$$g = \mathrm{diag}(0.99 \pm 0.02, -1.01 \pm 0.01, -1.01 \pm 0.02, -0.99 \pm 0.02)$$

$\longrightarrow$ Minkowski metric learnt!

trainable

Established qg Correlators

P. Komiske, E. Metodiev, J. Thaler arXiv:1810.05165 [hep-ph]

Extend CoLa+LoLa to tag q vs g:

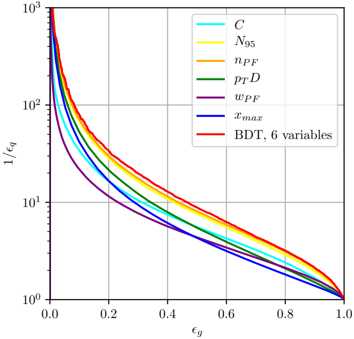
$$n_{\rm PF} = \sum_i 1 \qquad\qquad w_{\rm PF} = \frac{\sum_i p_{T,i}\Delta R_{i,\rm jet}}{\sum_i p_{T,i}}$$

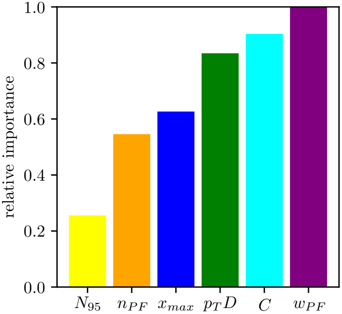
$$p_T D = \frac{\sqrt{\sum_i p_{T,i}^2}}{\sum_i p_{T,i}} \qquad\qquad C_{0.2} = \frac{\sum_{i,j} E_{T,i}E_{T,j}(\Delta R_{ij})^{0.2}}{\sum_i E_{T,i}^2}$$

$$x_{\rm max} = \max\left(\frac{p_{T,i}}{p_{T,\rm jet}}\right) \qquad\qquad N_{95} = \min(n), \;\; \text{where} \;\; \sum_i^n \frac{p_{T,i}}{p_{T,\rm jet}} \geq 0.95$$

$\longrightarrow$ IRC unsafe

Exploiting the Correlations





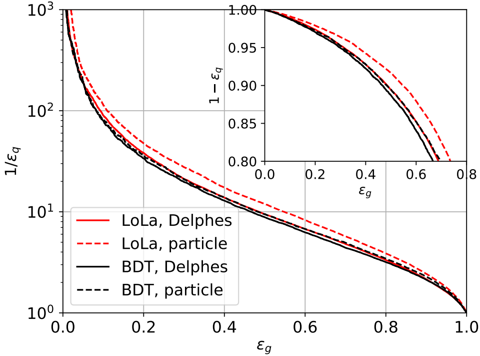
- Use BDT as a benchmark
- Important features not predictable

Include in LoLa

$$\tilde{k}_j \overset{\text{QG LoLa}}{\longrightarrow} \hat{k}'_j = \begin{pmatrix} m^2(\tilde{k}_j) \\ p_T(\tilde{k}_j) \\ \color{red}{p_T(\tilde{k}_j)\Delta R_{j,\text{jet}}} \\ w_{jm}^{(E)}\, E(\tilde{k}_m) \\ w_{jm}^{(d)}\, d_{jm}^2 \\ \color{red}{E_T(\tilde{k}_j)E_T(\tilde{k}_m)(\Delta R_{jm})^{0.2}} \end{pmatrix}$$

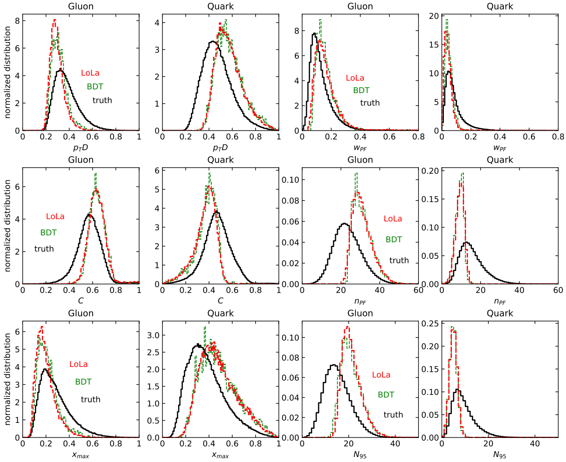
- Extend existing LoLa framework
- New observables encode qg relevant contributions

Understanding the Tagger



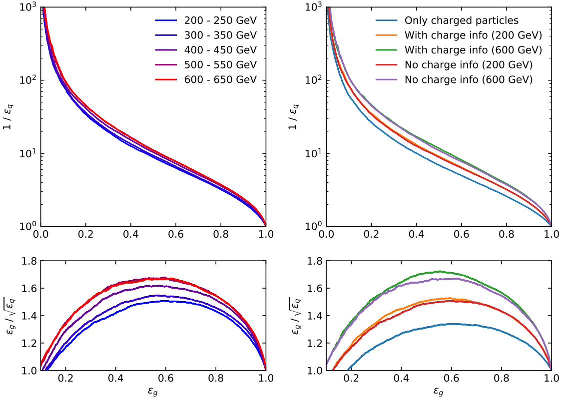
- LoLa-QG significantly outperforms BDT at particle level
- Still some improvement after Delphes detector simulation

Machine Learning on Pure Samples



$\longrightarrow$ NN and BDT learn similar features

Impact of Constituent Charge



- Can reproduce arXiv:1612.01551
- Include Delphes detector simulation
- *qg* separation improves with p_T

Stability in p_T

	Test			
$p_{T\text{jet}}$/GeV	200-250	300-350	400-450	600-650
200-250	0.813	0.818	0.805	0.74
300-350	0.811	0.825	0.823	0.80
400-450	0.809	0.824	0.834	0.80
600-650	0.807	0.816	0.830	0.841

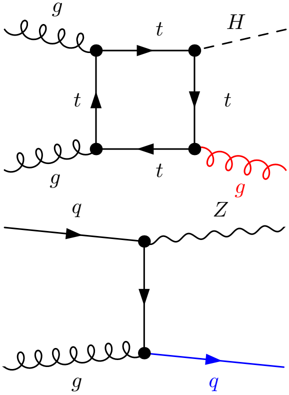
Train

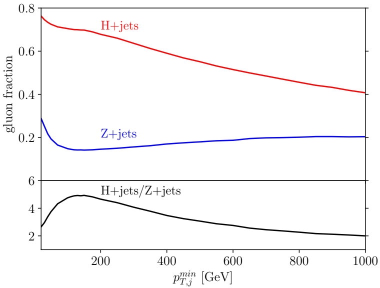
	200-210	210-220	220-230	230-240	240-250
200-210	0.812	0.812	0.812	0.818	0.816
210-220	0.812	0.813	0.812	0.819	0.817
220-230	0.804	0.805	0.810	0.811	0.808
230-240	0.803	0.804	0.801	0.814	0.809
240-250	0.810	0.811	0.811	0.820	0.818

- AUC stable across 50 GeV p_T slice
- Improvement at high p_T
- Uncertainty 1-2 on last digit

First Application: Monojet

Physics application





qg tagging did not improve WBF

A. Biekötter, F. Keilbach, R. Moutafis, T. Plehn, **J. Thompson** arXiv:1712.03973 [hep-ph]

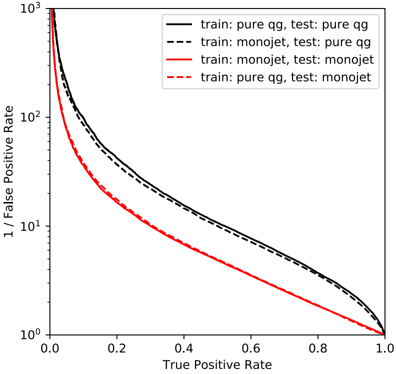
- Signal is gluon dominated, background quark dominated
- Purity decreases with increasing p_T

Tagging for Monojet

Train / Test: $p_{T\,\mathrm{jet}}$/GeV	200-250	250-300	300-350	600-650
200-250	0.691	0.683	0.674	0.604
250-300	0.691	0.685	0.677	0.605
300 - 350	0.687	0.683	0.677	0.614
600-650	0.630	0.638	0.646	0.631

- *qg* discrimination is good, but still lose in monojet
- More quarks in signal at high p_T
- impure signal plays off against increased tagging efficiency

Mixed Training-Testing



- Similar performance training on monojet/qg
- Drop in performance is not from qg tagging

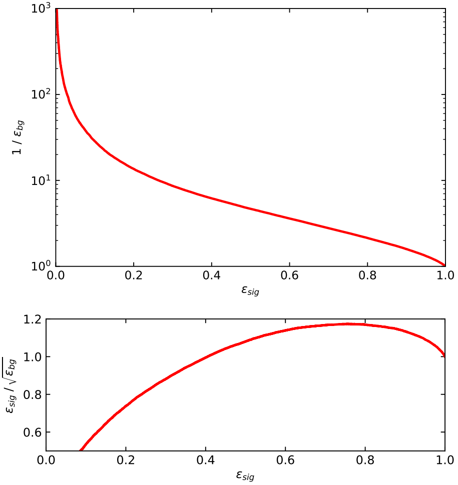
Second Application: Z' Searches

$$\mathcal{L} \supset g_q \sum_q Z' \bar{q} \gamma^\mu \gamma^5 q$$

- Democratic coupling to quarks
  $\longrightarrow$ Signal is quark dominated
  $\longrightarrow$ Background is gluon dominated
- Initial parameters:
  $\longrightarrow$ $M_{Z'} = 450$ GeV
  $\longrightarrow$ $g_q = 0.1$
- Can we increase the significance?

Observable	cut
N_j	≥ 2
p_{Tj1}^{min}	200GeV
p_{Tj1}^{max}	250GeV
p_{Tj2}	> 85GeV
$\lvert\eta_j\rvert$	< 2.8
$\lvert y^*\rvert$	< 0.3

First Z' results



- Consider the hardest jet
  $\longrightarrow$ Hardest 2 jets?
- 450 GeV Z'
  $\longrightarrow$ large QCD background
  $\longrightarrow$ Problem for triggering
- qg discrimination
  $\longrightarrow$ improve $S/\mathrm{sqrt}(b)$

Conclusions

- Detector simulation
  - Still possible to do qg tagging
  - Degrades performance
- Systematics from bin migration under control
- 2 physics applications for qg tagging
- Paper to come soon